

AUTOMATIC SEGMENTATION OF SKIN LESIONS USING HISTOGRAM THRESHOLDING

¹Nidhal Khdhair El Abbadi and ²Abbas Hussien Miry

¹Department of Computer Science, Education College, University of Kufa, Najaf, Iraq

²Department of Electrical Engineering, Engineering College, AL-Mustansiriyah University, Baghdad, Iraq

Received 2013-09-28; Revised 2013-10-02; Accepted 2013-12-19

ABSTRACT

Automatic lesion segmentation is an important part of computer-based image analysis of pigmented skin lesions. Currently, there is a great interest in the development of Computer-Aided Diagnosis (CAD) systems for dermoscopic images. The segmentation step is one of the most important ones, since its accuracy determines the eventual success or failure of a CAD system. This study introduced new method of dermoscopic images segmentation. The preprocess was the filtering operation to dermoscopy image to remove most of difficulties facing the efficient segmentations, like a variety of lesion shapes, sizes, color, changes due to different skin types and textures and presence of hairs. Segmentation based mainly on histogram thresholding. The enhancements of image achieved by using mathematical morphology in order to obtain better segmentation with smooth border and without any noise in the lesion region. The proposed method evaluated by using Hammoude Distance (HM) and the True Detection Rate (TDR). Also the proposed method is compared with other skin lesions segmentation methods such as Otsu, adaptive thresholding and fuzzy C-means. The accuracy of proposed method was 96.32%, which is highly promised result and dependable.

Keywords: Dermoscopy, Images Segmentation, Skin Lesions, Threshold, Histogram

1. INTRODUCTION

Malignant melanoma is the most common type of skin cancer; its spread has been growing rapidly over the past few decades. However, it is also the most treatable type of skin cancer if it is detected at an early stage (Silveira *et al.*, 2009). Dermoscopy (dermatoscopy or skin surface microscopy) is a non-invasive diagnostic technique for vivo observation of pigmented skin lesions used in dermatology.

This diagnostic tool allows for better perception of the structures of surface and subsurface and allows the recognition of the formal structures which are not visible to the naked eye and opens a new dimension of clinical morphological features of pigmented skin lesions. Dermoscopy and telemedicine have been important developments in the past few years. This will help improving the clinical diagnosis of skin lesions. Currently, there are many researchs to develop automated analysis for the images dermatoscopic which led to increase the chances

of early screening for teledermatology. There are three common steps in dermoscopic image analysis, which are: (1) Divide the image; (2) Extract features and features analysis and (3) Lesion classification. The most important stage is the segmentation due to its affects on the accuracy of the subsequent steps. Lesion segmentation is not easy process, due to: (1) There is low contrast between the lesion and the surrounding skin. (2) The border between skin and lesion is almost irregular. (3) Some of lesion mat contain some impurities, such as air bubble, hair or may be different skin textures. (4) Color diversity may be exhibit from the interior of the lesion (Xie and Bovik, 2013).

In computerized diagnosis of skin lesions, there are many suggested algorithms to segment the lesions. These algorithms try to isolate the lesion area from the rest of the image in order to permit an easier observation and diagnosis of the lesion. Some of the related works are.

Nourmohamadi and Pourghassem (2012) introduced new algorithm develop from intial segmentation by suggested a weighted combination for many

Corresponding Author: Nidhal Khdhair El Abbadi, Department of Computer Science, Education College, University of Kufa, Najaf, Iraq

segmentation algorithms like Fuzzy C-means, Otsu, K-means and Iterative algorithm. In addition, it will be possible to estimate the controlling parameters of the level set evolution from the obtained initial contour.

Wang *et al.* (2010) proposed new segmentation method. The first step in this method is the removing of hair and black board. The flooding variant of the watershed segmentation algorithm was implemented with features adapted to this domain. He suggested to use variance functions derived from the projection functions to add box to the lesion area which is estimated by using linear function to reduce the error in area estimation. The watershed border will be smoothing later by using B-splining smoothing method.

Xie and Bovik (2013) proposed new segmentation method for dermoscopy image based on combining two algorithms, Self Generating Neural Network and the Genetic Algorithm (GA). Seeds for GA is selected and used as initial neuron trees, the remaining samples training to generating Neural Forest. The second step is to cluster the image content and specify the number of clusters. These clusters can be combined to two region one for lesion and the other for background. This algorithm tested and compared with other segmentation algorithms.

1.1. K-Means Clustering Algorithm

K-means is learning method, easy to implementation on set of data to classify its to a number of clusters suggested in previous (k cluster). k-means solve the clustering problem by simply define number of centroids (k), each one will be generate cluster.

The algorithm assumes that the data features form a vector space and trying to find a compilation of the natural (Siddiqui and Mat Isa, 2012).

According to the K-means algorithm basic assembly and groups rely entirely on the choice of the midpoint kits. Euclidean distance used to determine the distances of the data elements suggested as K initial centers. The data elements assigned to the proper cluster according to the distance from the data elements to the centroids. The process is continued until no more changes occur in groups (Vora and Oza, 2013).

1.2. Fuzzy C-Means Clustering Algorithm

Fuzzy C-Means (FCM) is the commonly models used in clustering set of data. In FCM algorithm the distance from each point to the cluster center will be determined and according to that distance the members will be assigns, it is inverse relation. In this method the single data point may be assign to more than one group with specific membership grade. Often this method is used in pattern recognition.

FCM groups usually perform better than K-means compilation of overlapping blocks and noisy data. However, this algorithm is also stuck in the local Optima. It is known that both the K-means and FCM algorithms to be sensitive to outliers (Poddar and Mukhopadhyay, 2012).

1.3. Otsu Method

Otsu method is one of the oldest methods in the image segmentation. It is treated as a statistical method according to probabilistic implemented. It should be noted that the method Otsu is one of the best ways automatic threshold level. The basic principle in Otsu method is divided into two categories, such as foreground and the background. The threshold is obtained by finding automatic maximum contrast between the layers (Jassim and Altaani, 2013).

Otsu based on the pixel density contrast. Otsu method in the threshold level iterating through all possible Limit values and calculate the behavior of the spreading the pixels levels around each side of the threshold, any pixel may fall either in the foreground or background (Kaur and Kaur, 2011).

1.4. Adaptive Thresholding

In this algorithm segmented image is produced as a binary image from the gray scale or colour image. Threshold is determined for each point in the image (pixel). Each pixel compare with threshold to set it as background if it was below threshold, other wise set as foreground.

In Global adaptive thresholding the pixel intensity of each region have the role to grouping the foreground and background objects in separate region. Global thresholding almost depend on the image histogram to determine fixed threshold for all the pixels in the image, that mean it works when the histogram contains neatly separated peaks corresponding to the desired object(s) and background(s). Hence, it cannot deal with all images.

Local adaptive thresholding have different way to segment the image depending on finding threshold for each pixel according to the intensity in the neighbored pixels. Therefore, the local method can use to find threshold for images not thresholding with global.

Therefore, global and local thresholding sometimes used by researchers as a hybrid method.

Niblack and Sauvola use local variation technique. In these methods, the minimum is calculated based on the average local $M(X, Y)$ and δ standard deviation (X, Y) within the framework of size $W \times W$. Sauvola in the way is an improvement on the method Niblack, especially for colored documents and strongly illuminated (Singh *et al.*, 2011).

1.5. Niblack Algorithm

Niblack algorithm calculates the threshold pixel wise by moving the window over the image gray level. The calculation of the threshold is based on the mean m and local standard deviation of each pixel in the window as given below Equation (1):

$$T_{\text{Niblack}} = m + \sqrt{\frac{\sum p_i^2}{NP} - m^2} \quad (1)$$

where, NP is the number of pixels in the gray image, m is the average of pixel value p_i and K is fixed to-0.2 as suggested by (Saini and Dutta, 2012).

1.6. Sauvola's Algorithm

The Sauvola main idea depend on using the mean as indicator for the threshold. the threshold may be close to the mean when there are data with alot of local contrast, in the other side the threshold will be below the mean when there are littel contrast.

When using "integral image" accumulators for both mean and mean-squared pixel values the Sauvola will be more efficient. Later floating point arrays with 64 bits needed, this will be cost more for large images. Consequently, the tiled version will be gaved which have identical results of non tiled version. For each tiled version an accumulator arrays assign separatly in the memory.

1.7. Mathematical Morphology

Mathematical morphology is a branch of the non-linear signal processing fields and the application of the concepts of group theory to analyze the images. Conformation refers to the study of shapes and structures from a scientific perspective. Morphological filters or operators are non-linear transformations, which modify the geometric features of the images. These operators convert the input image into another image through repetition with another image in the form of a certain size which is known as structuring element. Structuring element is the infrastructure that will be used to analyze locally, for comparison, the shapes of interest. Choose the shape and size of the structuring element is very important: it affects the outcome; it may take the form of a disk of radius R , rectangular or other shape. There are many tools morphological image segmentation is corrosion, expansion, reconstruction (Aliiaout and Nasri, 2012).

1.8. Proposed Method

The proposed methods suggested three steps to achieve the desire result of segmented skin lesions and

before we do any one of these steps the color image will be converted to gray scale image and then we apply the following steps.

1.9. Filtering

The goal of the Wiener filter is to filter out noise that has corrupted a signal. It is based on a statistical approach and a more statistical account of the theory is given in the minimum Mean Square Errors (MMSE) estimator article.

The inverse filtering is a restoration technique for deconvolution, i.e., If we use low pass filter to bluer image, then we can restore image by inverse the filtering. However, inverse filtering is very sensitive to additive noise. Each type of degradation will be reduced separately (one each time) and this process can help us to develop a recover method for each degradation and simply gathering them. The Wiener filtering executes an optimal tradeoff between inverse filtering and noise smoothing. It removes the additive noise and inverts the blurring simultaneously.

The wiener filtering estimate the original image linearly and it is work perfectly in terms of mean square error. That means, the process of using the inverse filtering and noise smoothing reduce the entire mean square error to the minimum.

1.10. Histogram Segmentation

Practically speaking, let $H = [1, L]$ is the histogram of grayscale levels of image $f(x,y)$. The number of pixels with gray level i is denoted h_i . we choose the initial threshold θ (almost is at the mid of histogram) that divides the histogram range into two classes which are $H_1 = [1, \dots, \theta]$ and $H_2 = [\theta+1, \dots, L]$. Then, the gray level histogram for the two classes is Equation (2):

$$H_1 = [h_1, h_2, \dots, h_\theta], H_2 = [h_{\theta+1}, h_{\theta+2}, \dots, h_L] \quad (2)$$

Suppose the median for classes H_1 and H_2 are m_1, m_2 . Then, Let m^T be the overall median of the whole image (sum of m_1 and m_2).

In the next step, the new threshold θ is equal to m^T and according to it we divided the resulted image to two new groups H_1 and H_2 . This process will be repeated until no change in value of m^T . The main goal of this step is to binaries the gray scale image according to threshold θ as shown in **Fig. (1-3)**. The resulted image has two regions one for foreground (which represent the lesion part of image), while the other is the background which mainly represent the rest of skins.



Fig. 1. Original image with lesion



Fig. 4. Segmented image obtained by a proposed method after morphological



Fig. 2. Segmented image by expert medical

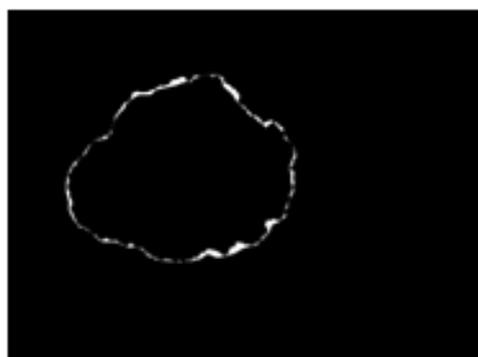


Fig. 5. Difference image between expert and proposed method



Fig. 3. Segmented image obtained by histogram

1.11. Mathematical Morphology

This step focuses on the enhancement the resulted image. We start to remove all the black parts inside the white region (some of them result from the dark hair) as shown in **Fig. 3** by using a morphological structuring element and smoothing the boundary edge between the two regions, the result image from this step shown in **Fig. 4**.

1.12. Evaluation

Two methods are applied to evaluate segmentation results for accuracy. Hammoude Distance (HM) and the True Detection Rate (TDR).

To implementing these metrics, let denote X to a resulted image from proposed automatic segmentation method and Y denote the segmented image achieved by expert physicians. Both X and Y is binary images.

There are two errors may be counted in this process, the first one is pixels classified as a pest by the medical expert while did not classified as pest by proposed method, where the second error is pixels classified as pest by proposed algorithm and did not classified as pest by a medical expert. **Figure 5** shows the difference between the segmented image by expert and proposed algorithm (white circle represent some of errors).

However, from a clinical point of view, the first type of error is more important since the pixels lesion should never be lost. Therefore, we felt the need to use two separate measures to take into account these two types of error.

1.13. Hammoude Distance (HM)

This distance measure used to measure the errors between the segmented part of physicians and the segmented part of proposed algorithm.

HM is based on a pixel by pixel comparison of the pixels enclosed by two boundaries Equation (3):

$$HM(X, Y) = \frac{\neq(X \cup Y) - \neq(X \cap Y)}{\neq(X \cup Y)} \quad (3)$$

1.14. True Detection Rate (TDR)

TDR metric measures the rate of pixels classified as lesion by both the automatic and the medical expert segmentation as Equation (4):

$$TDR(X, Y) = \frac{\neq(X \cap Y)}{\neq(Y)} \quad (4)$$

Table 1 shows the average scores for each method hash. The best way according to HM is proposed with 11.37%. **Figure 6-8** shows the segmentation of different methods.

Table 1. Results of test images for different methods

Methods	HM (%)	TDR (%)
K-means	13.87	93.66
Fuzzy K-means	13.07	93.83
OTSU	13.89	93.01
Niblack	16.73	83.15
Sauvola	16.45	89.90
Proposed	11.37	96.32

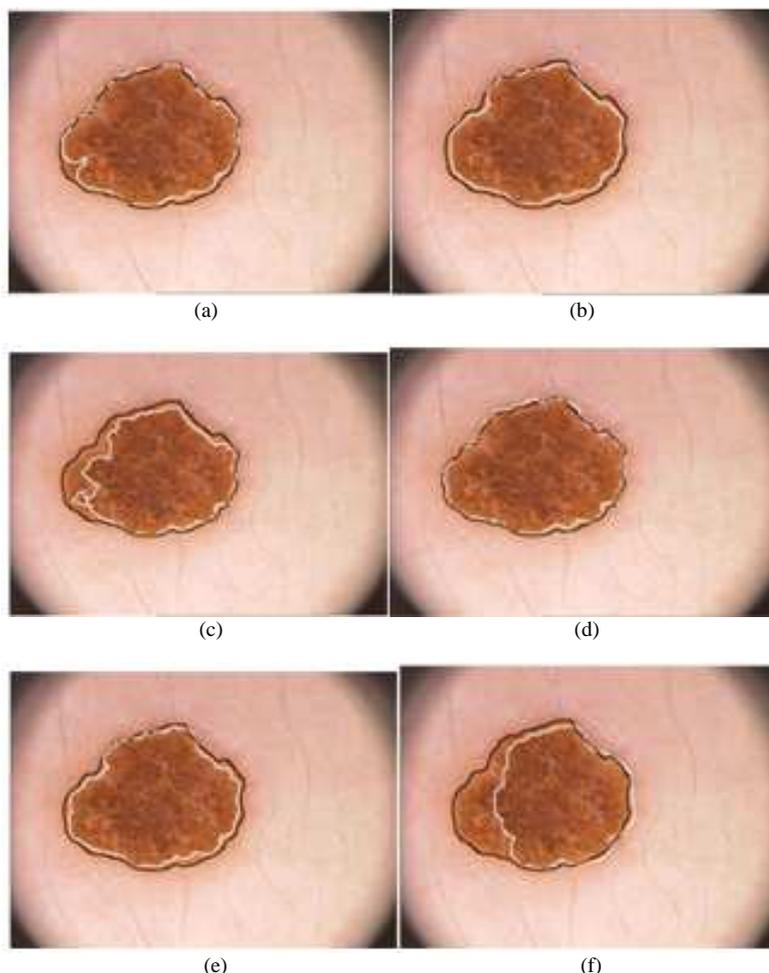


Fig. 6. Segmentation of the dermatoscopic image by: Black border by (expert segmentation), white border with (six approaches, (a) k-mean, (b) FCM (c) Otsu (d) Niblack, (e) Sauvola, (f) Proposed method

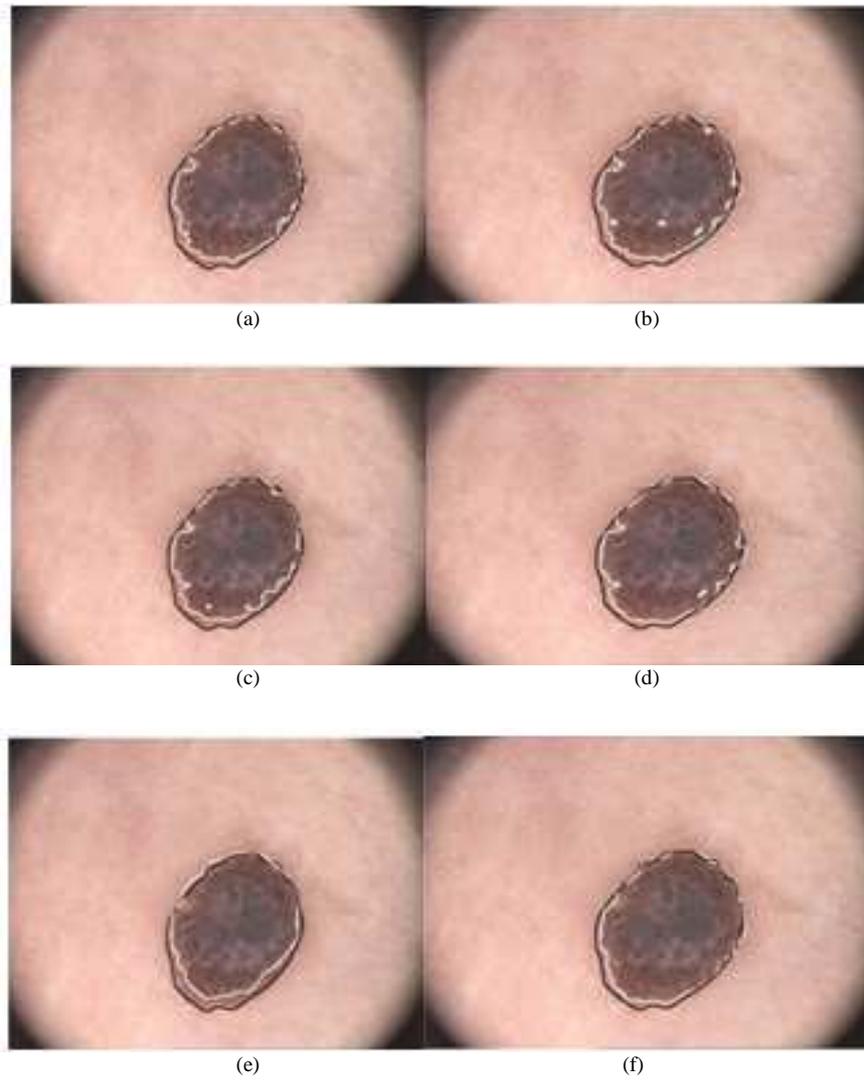


Fig. 7. Segmentation of the dermatoscopic image by: black border (expert segmentation), white border (six approaches), (a) k-mean, (b) FCM, (c) Otsu, (d) Niblack, (e) Sauvola, (f) Proposed method



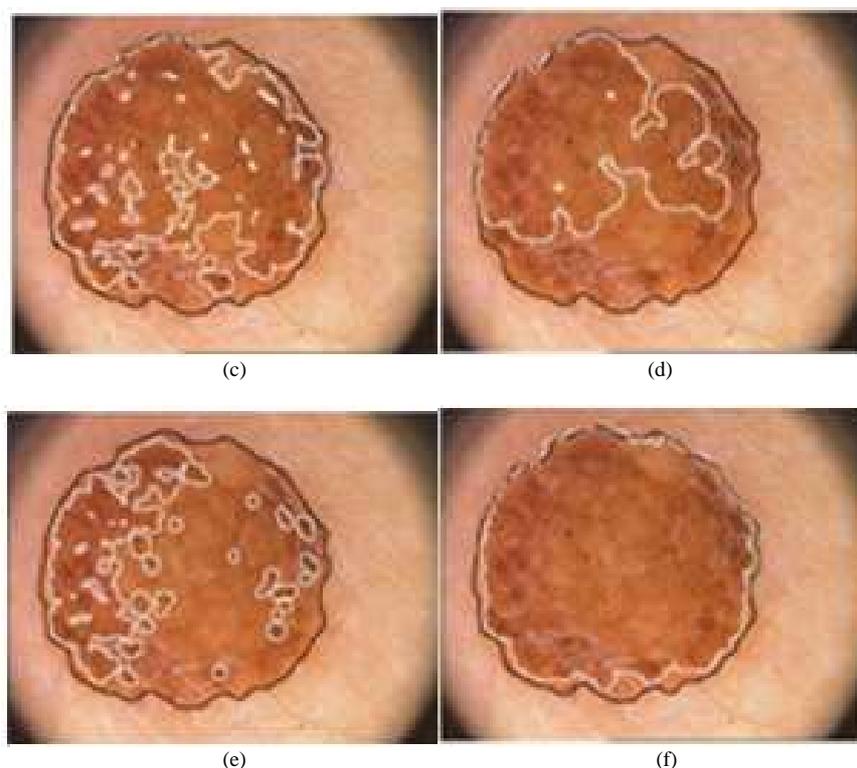


Fig. 8. Segmentation of the dermatoscopic image by: black border (expert segmentation), white border (six approaches) (a) k-mean, (b) FCM, (c) Otsu, (d) Niblack, (e) Sauvola, (f) Proposed method

2. CONCLUSION

In this study, we introduced new segmentation method depending on histogram thresholding. The proposed algorithm removes the effects of hair and other noise by using Wiener filter. Proposed algorithm evaluated by comparing its segmented images with images segmented by experts medical (**Fig. 1 and 2**) and measuring the distance between them by using two metrics (HM and TDR). Also the proposed algorithm compared with other technique and gives highly promised results as shown in **Table 1**. This approach can be used efficiently to support decision-making problems in skin diagnosis. We believe that proposed methods are useful for segmentation of the pest in aided diagnostic computer system to assist the clinical diagnosis of dermatologists.

3. REFERENCES

Aliiaout, A. and M. Nasri, 2012. Medical image segmentation by marker controlled watershed and mathematical morphology. *Int. J. Multimedia Applic.*, 4: 1-9. DOI: 10.5121/ijma.2012.4301

- Jassim, F.A. and F.H. Altaani, 2013. Hybridization of OTSU method and median filter for color image segmentation. *Int. J. Soft Comput. Eng.*, 3: 69-74.
- Kaur, N. and R. Kaur, 2011. A review on various methods of image thresholding. *Int. J. Comput. Sci. Eng.*, 3: 3441-3443.
- Nourmohamadi, M. and H. Pourghassem, 2012. Dermoscopy image segmentation using a modified level set algorithm. *Proceedings of the 4th International Conference on Computational Intelligence and Communication Networks*, Nov. 3-5, IEEE Xplore Press, Mathura, pp: 286-290. DOI: 10.1109/CICN.2012.80
- Poddar, S. and A. Mukhopadhyay, 2012. Cluster: A MATLAB GUI package for data clustering. *Int. Technol. Res. Lett.*, 1: 78-83.
- Saini, R. and M. Dutta, 2012. Image segmentation for uneven lighting images using adaptive thresholding and dynamic window based on incremental window growing approach. *Int. J. Comput. Applic.*, 56: 31-36. DOI: 10.5120/8954-3140

- Siddiqui, F.U. and N.A. Mat Isa, 2012. Optimized K-Means (OKM) clustering algorithm for image segmentation. *Opto-Electron. Rev. J.*, 20: 216-225. DOI: 10.2478/s11772-012-0028-8
- Silveira, M., J.C. Nascimento, J.S. Marques, A.R.S. Marcal and T. Mendonca *et al.*, 2009. Comparison of segmentation methods for melanoma diagnosis in dermoscopy images. *IEEE J. Selected Top. Signal Process.*, 3: 35-45. DOI: 10.1109/JSTSP.2008.2011119
- Singh, T.R., S. Roy, O.I. Singh, T. Sinam and K.M. Singh, 2011. A new local adaptive thresholding technique in binarization. *Int. J. Comput. Sci.*, 8: 271-277.
- Vora, P. and B. Oza, 2013. A survey on k-mean clustering and particle swarm optimization. *Int. J. Sci. Modern Eng.*, 1: 24-26.
- Wang, H., X. Chen, R.H. Moss, R.J. Stanley and W.V. Stoecker *et al.*, 2010. Watershed segmentation of dermoscopy images using a watershed technique. *Skin Res. Technol. J.*, 16: 378-384. DOI: 10.1111/j.1600-0846.2010.00445.x.
- Xie, F. and A.C. Bovik, 2013. Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm. *Patt. Recogn.*, 46: 1012-1019. DOI: 10.1016/j.patcog.2012.08.012